



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/676,299	09/30/2003	Lance G. Laing	04107/100L443-US3	8447

7278 7590 07/13/2004

DARBY & DARBY P.C.

P. O. BOX 5257

NEW YORK, NY 10150-5257

EXAMINER

DUNSTON, JENNIFER ANN

ART UNIT

PAPER NUMBER

1636

DATE MAILED: 07/13/2004

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary

Application No.

10/676,299

Applicant(s)

LAING, LANCE G.

Examiner

Jennifer Dunston

Art Unit

1636

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --
Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

- 1) ☒ Responsive to communication(s) filed on 23 February 2004.
- 2a) ☐ This action is **FINAL**. 2b) ☒ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

- 4) ☒ Claim(s) 1-4 is/are pending in the application.
- 4a) Of the above claim(s) _____ is/are withdrawn from consideration.
- 5) ☐ Claim(s) _____ is/are allowed.
- 6) ☒ Claim(s) 1-4 is/are rejected.
- 7) ☐ Claim(s) _____ is/are objected to.
- 8) ☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

- 9) ☐ The specification is objected to by the Examiner.
- 10) ☐ The drawing(s) filed on _____ is/are: a) ☐ accepted or b) ☐ objected to by the Examiner.
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

- 12) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☐ All b) ☐ Some * c) ☐ None of:
1. ☐ Certified copies of the priority documents have been received.
2. ☐ Certified copies of the priority documents have been received in Application No. _____.
3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

- 1) ☒ Notice of References Cited (PTO-892)
- 2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
- 3) ☒ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08)
Paper No(s)/Mail Date 11/26/2003.
- 4) ☐ Interview Summary (PTO-413)
Paper No(s)/Mail Date. _____.
- 5) ☐ Notice of Informal Patent Application (PTO-152)
- 6) ☒ Other: Sequence Search Attachments.

DETAILED ACTION

Claims 1-4 are pending in the instant application.

Information Disclosure Statement

Receipt of an information disclosure statement, filed on 11/23/2003, is acknowledged.

The signed and initialed PTO 1449 has been mailed with this action.

Claim Objections

Claim 4 is objected to because of the following informalities: the claim is grammatically incorrect in that the claim does not clearly indicate the numbers 4, 6, 8 and 10 are sequence identifiers. It would be remedial to amend the claim language to read something like "SEQ ID NOS: 3 and 4; SEQ ID NOS: 5 and 6."

Claim Rejections - 35 USC § 112

The following is a quotation of the second paragraph of 35 U.S.C. 112:

The specification shall conclude with one or more claims particularly pointing out and distinctly claiming the subject matter which the applicant regards as his invention.

Claims 1-4 are rejected under 35 U.S.C. 112, second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter which applicant regards as the invention.

Claim 1 is indefinite in that the metes and bounds of the term "have" are unclear. The term "have" can be interpreted as "consisting of" or "comprising". It would be remedial to amend the claim to use either open or closed language.

Claim 2 is indefinite in that the metes and bounds of the term “has” are unclear. The term “has” can be interpreted as “consisting of” or “comprising”. It would be remedial to amend the claim to use either open or closed language.

Claim 4 is indefinite in that the metes and bounds of the term “comprised of” are unclear. It would be remedial to amend the claim to use “consisting of”, “comprises”, or “comprising.”

Claim Rejections - 35 USC § 102

The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

A person shall be entitled to a patent unless –

(b) the invention was patented or described in a printed publication in this or a foreign country or in public use or on sale in this country, more than one year prior to the date of application for patent in the United States.

Claims 1-4 are drawn to oligonucleotides comprising a sequence that differs by no more than three bases or base pairs from a sequence selected from the group consisting of SEQ ID NOS: 3-10, wherein the oligonucleotides may be double stranded and may be hybrid pairs selected from the group consisting of SEQ ID NOS: 3 and 4, SEQ ID NOS: 5 and 6, SEQ ID NOS: 7 and 8 and SEQ ID NOS: 9 and 10. The specification defines an oligonucleotide as a nucleic acid, generally of at least 10, preferably at least 15, and more preferably at least 20 nucleotides, preferably no more than 100 nucleotides, that contains a specific protein binding site (page 10, lines 6-9). Given the broadest reasonable interpretation, the claimed invention encompasses any oligonucleotide comprising a sequence selected from the group of SEQ ID NOS: 3-10.

Claims 1-4 are rejected under 35 U.S.C. 102(b) as being anticipated by Wu et al (The Journal of Biological Chemistry, Vol. 268, No. 1, pages 52-58, 1993; see the entire reference) as evidenced by San Francisco et al (Nucleic Acids Research, Vol. 18, No. 3, pages 619-624, 1990; see the entire reference).

Wu et al teach a 153 bp DNA fragment consisting of the *ars* promoter of *E. coli* plasmid R773 from nucleotides -105 to +48 (e.g. page 52, Preparation of DNA Fragments for the Gel Retardation and Footprint Assays; page 53, Promoter Region). Wu et al teach that the 153 bp DNA fragment contains the ArsR binding site (e.g. Figure 9). SEQ ID NOS: 3, 4, 7 and 8 are 100% identical to the nucleotide sequence shown within Figure 9. The sequence shown within Figure 9 is a portion of the sequence contained within plasmid pJHW1, which consists of a 0.73 kb EcoRI-HindIII fragment of pWSU1 cloned into plasmid pJBS633 (e.g. Table 1).

San Francisco et al teach plasmid pWSU1 and the 0.73 kb EcoRI-HindIII fragment of pWSU1 (e.g. page 619, Bacterial strains, plasmids and bacteriophage). San Francisco et al cloned the 0.73 kb EcoRI-HindIII fragment into plasmids pUC18 and pUC19 and phages M13mp8 and M13mp9 for sequencing (e.g. page 619 Bacterial strains, plasmids and bacteriophage). The sequence of the 0.73 kb EcoRI-HindIII fragment is contained in GenBank Accession No. X16045. The nucleic acids from bases 42-71, 71-44, 42-64 and 64-44 are 100% identical to SEQ ID NOS: 3, 4, 7 and 8 (alignments provided), respectively. Further, as indicated in the attached alignments, SEQ ID NOS: 3 and 4 and SEQ ID NOS: 7 and 8 are overlapping within GenBank Accession No. X16045.

Therefore, Wu et al necessarily teach oligonucleotides comprising SEQ ID NOS: 3, 4, 7 and 8 and hybrid pairs of SEQ ID NOS: 3 and 4 and SEQ ID NOS: 7 and 8.

Claims 1-4 are rejected under 35 U.S.C. 102(b) as being anticipated by Xu et al (The Journal of Biological Chemistry, Vol. 271, No. 5, pages 2427-2432, 1996; see the entire reference) as evidenced by Diorio et al (Journal of Bacteriology, Vol. 177, No. 8, pages 2050-2056, 1995; see the entire reference).

Xu et al teach a 208 bp fragment containing the *E. coli* chromosomal *arsR* operator (page 2429, Gel Mobility Shift and DNaseI Footprinting Assays). Xu et al generated this fragment using the forward primer 5'-CGGAATTCCGACGCAAAGTC-3' and reverse primer 5'-CCAGACGGGTTTCATCAGCAAGAATTTTG-3' for polymerase chain reaction amplification of the *arsR* operator region (e.g. page 2429, Primer Extension; page 2429, Gel Mobility Shift and DNaseI Footprinting Assays).

Diorio et al teach the sequence of the *E. coli* chromosomal *ars* operon (e.g. Figure 2 and GenBank Accession No. X80057). The primer sequences used by Xu et al are contained in the *E. coli* chromosomal *ars* operon sequence of GenBank Accession No. X80057. The underlined sequence of the forward primer corresponds to bases 534-546 of X80057. The reverse primer corresponds to bases 765-737 of X80057. Further, the nucleic acid sequences from bases 621-662, 662-623, 638-662 and 662-638 of X80057 are 100% identical to SEQ ID NOS: 5, 6, 9 and 10 (alignment provided), respectively. Moreover, as indicated in the attached alignments, SEQ ID NOS: 5 and 6 and SEQ ID NOS: 9 and 10 are overlapping within GenBank Accession No. X80057.

Therefore, Xu et al necessarily teach oligonucleotides comprising SEQ ID NOS: 5, 6, 9 and 10 and hybrid pairs SEQ ID NOS: 5 and 6 and SEQ ID NOS: 9 and 10.

Conclusion

No claims are allowed.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Jennifer Dunston whose telephone number is 571-272-2916. The examiner can normally be reached on M-F, 9 am to 5 pm.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Remy Yucel can be reached on 571-272-0781. The fax phone number for the organization where this application or proceeding is assigned is 703-872-9306.

Any inquiry of a general nature or relating to the status of this application or proceeding should be directed to (571) 272-0547.

Patent applicants with problems or questions regarding electronic images that can be viewed in the Patent Application Information Retrieval system (PAIR, <http://pair-direct.uspto.gov>) can now contact the USPTO's Patent Electronic Business Center (Patent EBC) for assistance. Representatives are available to answer your questions daily from 6 am to midnight (EST). The toll free number is (866) 217-9197. When calling please have your application serial or patent number, the type of document you are having an image problem with, the number of pages and the specific nature of the problem. The Patent Electronic Business Center will notify applicants of the resolution of the problem within 5-7 business days. Applicants can also check PAIR to confirm that the problem has been corrected. The USPTO's Patent Electronic Business Center is a complete service center supporting all patent business on the Internet. The USPTO's PAIR system provides Internet-based access to patent application status and history information. It also enables applicants to view the scanned images of their own application file folder(s) as well as general patent information available to the public.

For all other customer support, please call the USPTO Call Center (UCC) at 800-786-9199.


GERRY LEFFERS
PRIMARY EXAMINER

Jennifer Dunston
Examiner
Art Unit 1636

jad

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 26, 2004, 15:20:44 ; Search time 417.308 Seconds
(without alignments)
3115.905 Million cell updates/sec

Title: US-10-676-299-3

Perfect score: 30

Sequence: 1 ttaatacatatgcgtttttggttggttg 30

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.:

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg.hum.*
- 31: em.htg.inv.*
- 32: em.htg.other.*
- 33: em.htg.mus.*
- 34: em.htg.pln.*
- 35: em.htg.rod.*
- 36: em.htg.man.*
- 37: em.htg.vrt.*
- 38: em.sy.*
- 39: em.htgo.hum.*
- 40: em.htgo.mus.*
- 41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	30	100.0	727	1	ECRFARSR	X16045 E. coli R-f
2	28.4	94.7	120826	1	AP005147	AP005147 Salmonella
3	23.2	77.3	177109	2	AC118161	AC118161 Rattus no
4	23.2	77.3	199985	2	AC130151	AC130151 Rattus no
5	22.6	73.3	152772	9	AC095326	AC095326 Homo sapi
6	22	73.3	177672	2	AC010959	AC010959 Homo sapi
7	22	73.3	183748	9	AC124915	AC124915 Homo sapi
8	22	73.3	207683	2	AC108840	AC108840 Mus muscu
9	22	73.3	223489	5	EX248100	EX248100 Zebrafish
10	22	73.3	295094	2	EX649429	EX649429 Danio rer
11	21.6	72.0	10391	6	AX323523	AX323523 Sequence
12	21.6	72.0	10099	1	AE002489	AE002489 Neisseria
13	21.6	72.0	148153	9	AL513533	AL513533 Human DNA
14	21.6	72.0	169613	2	EX321886	EX321886 Danio rer
15	21.6	72.0	174303	10	AC144801	AC144801 Mus muscu
16	21.6	72.0	182051	9	AC007638	AC007638 Homo sapi
17	21.6	72.0	207265	2	AC122227	AC122227 Mus muscu
18	21.6	72.0	214815	2	AC130222	AC130222 Mus muscu
19	21.6	72.0	221911	2	AC114095	AC114095 Rattus no
20	21.6	72.0	241420	2	AC094501	AC094501 Rattus no
21	21.6	72.0	349980	6	AX044032	AX044032 Sequence
22	21.2	70.7	73000	2	AC090562	AC090562 Homo sapi
23	21.2	70.7	88747	9	AC104409	AC104409 Homo sapi
24	21.2	70.7	160811	9	AC020553	AC020553 Homo sapi
25	21.2	70.7	184897	9	AC090919	AC090919 Homo sapi
26	21	70.0	1465	5	BC049518	BC049518 Danio rer
27	21	70.0	2664	9	D89962	D89962 Homo sapien
28	21	70.0	4133	9	AB032158	AB032158 Homo sapi
29	21	70.0	65542	9	AC091766	AC091766 Homo sapi
30	21	70.0	110685	10	AL928957	AL928957 Mouse DNA
31	21	70.0	131289	2	AC147278	AC147278 Pan trogl
32	21	70.0	137156	2	AC138846	AC138846 Homo sapi
33	21	70.0	140410	2	AC139483	AC139483 Homo sapi
34	21	70.0	145692	9	AL355303	AL355303 Human DNA
35	21	70.0	148385	9	HSJ520B18	HSJ520B18 Human DNA
36	21	70.0	150959	2	EX322574	EX322574 Danio rer
37	21	70.0	151259	2	AC138822	AC138822 Homo sapi
38	21	70.0	154803	9	AC131392	AC131392 Homo sapi
39	21	70.0	157792	9	AC069209	AC069209 Homo sapi
40	21	70.0	158755	2	AC145032	AC145032 Homo sapi
41	21	70.0	160127	2	AC068977	AC068977 Homo sapi
42	21	70.0	160701	9	AC108106	AC108106 Homo sapi
43	21	70.0	160901	2	AC138812	AC138812 Homo sapi
44	21	70.0	161000	2	AC145100	AC145100 Homo sapi
45	21	70.0	161362	2	AL355503	AL355503 Homo sapi

ALIGNMENTS

RESULT 1	ECRFARSR	727 bp	DNA	linear	BCT 07-SEP-1994
LOCUS	E. coli R-factor R773 arsR gene.				
DEFINITION	X16045				
ACCESSION	X16045.1				
VERSION	GI:42716				
KEYWORDS	antibiotic resistance; arsR gene; ArsR protein; DNA-binding protein; regulatory protein; resistance gene.				
SOURCE	Escherichia coli				
ORGANISM	Escherichia coli				
REFERENCE	1 (bases 1 to 727)				
AUTHORS	San Francisco, M.J., Hope, C.L., Owolabi, J.B., Tisa, L.S. and Rosen, B.P.				

TITLE Identification of the metalloregulatory element of the plasmid-encoded arsenical resistance operon
JOURNAL Nucleic Acids Res. 18 (3), 619-624 (1990)
MEDLINE 90174986
PUBMED 2408017
REFERENCE 2 (bases 1 to 727)
AUTHORS Rosen, B.P.
TITLE Direct Submission
JOURNAL Submitted (09-AUG-1989) Rosen B.P., Department of Biochemistry, Wayne State University, School of Medicine, 54- E Canfield Avenue, Detroit MI 48201, U S A

FEATURES
source Location/Qualifiers
1..727
/organism="Escherichia coli"
/mol_type="genomic DNA"
/db_xref="taxon:562"
/clone="pWSU1"
/clone_lib="pBR322"
73..79
/note="pot. -35 region"
96..102
/note="pot. -10 region"
107
/note="transcriptional start site"
114..118
/note="pot. ribosome binding site"
125..478
/note="unnamed protein product; ArsR protein (AA 1 - 117)"
/codon_start=1
/transl_table=1
/protein_id="CAA34168.1"
/db_xref="GI:42717"
/db_xref="GOA:P15905"
/db_xref="SWISS-PROT:P15905"
/translation="MLQTPQLQFLNLSDETRLGIVLLIREMGEICVDCMLDQSQPKSRHMLRESGILLDRKQKWHYRLSPWAAQIEQAWLSQDDVQVIARKLASVNCSSGSAVCI"
482..511
terminator
/note="pot. stem-loop structure"

ORIGIN
Query Match 100.0%; Score 30; DB 1; Length 727;
Best Local Similarity 100.0%; Pred. No. 0.4;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTAATCATATCGCTTTTGGTATGTGTTG 30
|||||
Db 42 TTAATCATATCGCTTTTGGTATGTGTTG 71

RESULT 2
AP005147
LOCUS AP005147 120826 bp DNA circular BCT 14-MAY-2002
DEFINITION Salmonella typhimurium plasmid R64 DNA, complete sequence.
ACCESSION AP005147
VERSION AP005147.1 GI:20521502
KEYWORDS
SOURCE Salmonella typhimurium
ORGANISM Salmonella typhimurium
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Salmonella.

REFERENCE 1
AUTHORS Komano, T., Kubo, A. and Nishioka, T.
TITLE Shufflon: Multi-inversion of four contiguous DNA segments of plasmid R64 creates seven different open reading frames
JOURNAL Nucleic Acids Res. 15 (3), 1165-1172 (1987)
MEDLINE 87146423
PUBMED 3029698
REFERENCE 2
AUTHORS Kubo, A., Kusukawa, A. and Komano, T.
TITLE Nucleotide sequence of the rci gene encoding shufflon-specific DNA recombinase in the IncII plasmid R64: homology to the site-specific recombinases of integrase family

JOURNAL Mol. Gen. Genet. 213 (1), 30-35 (1988)
MEDLINE 89137142
PUBMED 3065610
REFERENCE 3
AUTHORS Komano, T., Toyoshima, A., Morita, K. and Nishioka, T.
TITLE Cloning and nucleotide sequence of the orit region of the IncII plasmid R64
JOURNAL J. Bacteriol. 170 (9), 4385-4387 (1988)
MEDLINE 88314948
PUBMED 3045094
REFERENCE 4
AUTHORS Furuya, N., Nishioka, T. and Komano, T.
TITLE Nucleotide sequence and functions of the orit operon in IncII plasmid R64
JOURNAL J. Bacteriol. 173 (7), 2231-2237 (1991)
MEDLINE 9117811
PUBMED 1848811
REFERENCE 5
AUTHORS Furuya, N. and Komano, T.
TITLE Determination of the nick site at orit of IncII plasmid R64: Global similarity of orit structures of IncII and IncP plasmids
JOURNAL J. Bacteriol. 173 (20), 6612-6617 (1991)
MEDLINE 92011438
PUBMED 1917882
REFERENCE 6
AUTHORS Kim, S.R., Funayama, N. and Komano, T.
TITLE Nucleotide sequence and characterization of the traABCD region of IncII plasmid R64
JOURNAL J. Bacteriol. 175 (17), 5035-5042 (1993)
MEDLINE 93352408
PUBMED 8349545
REFERENCE 7
AUTHORS Furuya, N. and Komano, T.
TITLE Surface exclusion gene of IncII plasmid R64: nucleotide sequence and analysis of deletion mutants
JOURNAL Plasmid 32 (1), 80-84 (1994)
MEDLINE 95083745
PUBMED 7991676
REFERENCE 8
AUTHORS Furuya, N. and Komano, T.
TITLE Nucleotide sequence and characterization of the trbABC region of the IncII plasmid R64: existence of the pnd gene for plasmid maintenance within the transfer region
JOURNAL J. Bacteriol. 178 (6), 1491-1497 (1996)
MEDLINE 96198148
PUBMED 8626273
REFERENCE 9
AUTHORS Kim, S.R. and Komano, T.
TITLE The plasmid R64 thin pilus identified as a type IV pilus
JOURNAL J. Bacteriol. 179 (11), 3594-3603 (1997)
MEDLINE 97315231
PUBMED 9171405
REFERENCE 10
AUTHORS Narahara, K., Rahman, E., Furuya, N. and Komano, T.
TITLE Requirement of a limited segment of the sog gene for plasmid R64 conjugation
JOURNAL Plasmid 38 (1), 1-11 (1997)
MEDLINE 97428559
PUBMED 9281491
REFERENCE 11
AUTHORS Furuya, N. and Komano, T.
TITLE Mutational analysis of the R64 orit region: requirement for precise location of the Nika-binding sequence
JOURNAL J. Bacteriol. 179 (23), 7291-7297 (1997)
MEDLINE 98053841
PUBMED 9393692
REFERENCE 12
AUTHORS Yoshida, T., Furuya, N., Ishikura, M., Isobe, T., Haino-Fukushima, K., Ogawa, T. and Komano, T.
TITLE Purification and characterization of thin pili of IncII plasmids Colib-P9 and R64: formation of PiliV-specific cell aggregates by type IV pili
JOURNAL J. Bacteriol. 180 (11), 2842-2848 (1998)
MEDLINE 98053841
PUBMED 9393692

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 26, 2004, 15:20:44 ; Search time 399.487 Seconds
(without alignments)
3115.905 Million cell updates/sec

Title: US-10-676-299-4
Perfect score: 28
Sequence: 1 cacacataacacaaacgcatacgattgatt 28

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 5940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_em.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_on.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_rod.*
- 36: em_htg_nam.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
C 1	28	100.0	727	1	ECRFARSR	X16045 E. coli R-f
C 2	26.4	94.3	120826	1	AP005147	AP005147 Salmonell
C 3	21.6	77.1	169613	2	BX321886	BX321886 Dario rer
C 4	21.6	77.1	182051	9	AC007638	AC007638 Homo sapi
C 5	21.2	75.7	73000	2	AC090562	AC090562 Homo sapi
C 6	21.2	75.7	160811	9	AC020553	AC020553 Homo sapi
C 7	21.2	75.7	174303	10	AC144801	AC144801 Mus muscu
C 8	21.2	75.7	177109	2	AC118161	AC118161 Rattus no
C 9	21.2	75.7	184897	9	AC090919	AC090919 Homo sapi
C 10	21.2	75.7	199985	2	AC130151	AC130151 Rattus no
C 11	20.6	73.6	6391	6	AX323523	AX323523 Sequence
C 12	20.6	73.6	10099	1	AE002489	AE002489 Neisseria
C 13	20.6	73.6	15542	9	AC091766	AC091766 Homo sapi
C 14	20.6	73.6	116885	10	AL928957	AL928957 Mouse DNA
C 15	20.6	73.6	121836	2	AC122166	AC122166 Medicago
C 16	20.6	73.6	148151	9	AL513533	AL513533 Human DNA
C 17	20.6	73.6	160127	2	AC068977	AC068977 Homo sapi
C 18	20.6	73.6	166762	2	AC055852	AC055852 Homo sapi
C 19	20.6	73.6	198788	8	ATCHRIV46	AL161546 Arabidops
C 20	20.6	73.6	207674	8	AFEC88	Z97343 Arabidopsis
C 21	20.6	73.6	217584	2	AC113299	AC113299 Mus muscu
C 22	20.6	73.6	218580	2	AC113027	AC113027 Mus muscu
C 23	20.6	73.6	349980	6	AX044332	AX044332 Sequence
C 24	20.2	72.1	34796	3	CEP53C11	Z79756 Caenorhabdi
C 25	20.2	72.1	151930	9	CNS01DVT	AL136038 Human chr
C 26	20.2	72.1	176345	2	AC132913	AC132913 Mus muscu
C 27	20.2	72.1	180044	2	AC136679	AC136679 Mus muscu
C 28	20.2	72.1	181902	2	AC026286	AC026286 Homo sapi
C 29	20.2	72.1	185702	2	AC138246	AC138246 Mus muscu
C 30	20.2	72.1	224453	2	AC134283	AC134283 Rattus no
C 31	20.2	72.1	230082	2	AC110303	AC110303 Rattus no
C 32	20.2	72.1	245066	2	AC125734	AC125734 Rattus no
C 33	20	71.4	1553	6	AX078799	AX078799 Sequence
C 34	20	71.4	1553	6	BD013422	BD013422 Transcrip
C 35	20	71.4	7276	8	AB089813	AB089813 Dauscu ca
C 36	20	71.4	12787	1	AE015654	AE015654 Shewanell
C 37	20	71.4	88557	9	AL136314	AL136314 Human DNA
C 38	20	71.4	105320	3	AC115680	AC115680 Dictyoste
C 39	20	71.4	108582	2	AC120459	AC120459 Homo sapi
C 40	20	71.4	119716	2	AC125480	AC125480 Medicago
C 41	20	71.4	131289	2	AC147278	AC147278 Pan trogl
C 42	20	71.4	137156	2	AC138846	AC138846 Homo sapi
C 43	20	71.4	138005	2	AC119008	AC119008 Rattus no
C 44	20	71.4	140410	2	AC139483	AC139483 Homo sapi
C 45	20	71.4	149030	9	AC005375	AC005375 Homo sapi

ALIGNMENTS

RESULT 1
ECRFARSR/c
LOCUS E. coli R-factor R773 arsr gene.
DEFINITION X16045
ACCESSION X16045
VERSION GI:42716
KEYWORDS
arsenical resistance; arsr gene; Arsr protein; DNA-binding protein;
regulatory protein; resistance gene.
SOURCE
ORGANISM
Escherichia coli
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.
REFERENCE
1 (bases 1 to 727)
AUTHORS
San Francisco, M.J., Hope, C.L., Owlabi, J.B., Tisa, L.S. and
Rosen, B.P.

ECRFARSR 727 bp DNA linear ECT 07-SEP-1994

TITLE Identification of the metalloregulatory element of the plasmid-encoded arsenical resistance operon
JOURNAL Nucleic Acids Res. 18 (3), 619-624 (1990)
MEDLINE
PUBMED 90174986
REFERENCE 2408017
AUTHORS Rosen, B.P.
TITLE Direct Submission
JOURNAL Submitted (09-AUG-1989) Rosen B.P., Department of Biochemistry, Wayne State University, School of Medicine, 54- E Canfield Avenue, Detroit MI 48201, U S A

FEATURES
source
1..727
/organism="Escherichia coli"
/mol_type="genomic DNA"
/db_xref="taxon:562"
/clone="pWSU1"
/clone_lib="pBR322"
73..79
/note="pot. -35 region"
96..102
/note="pot. -10 region"
107
/note="transcriptional start site"
114..118
/note="pot. ribosome binding site"
125..478
/note="unnamed protein product; ArsR protein (AA 1 - 117)"
/codon_start=1
/transl_table=11
/protein_id="CA934168.1"
/db_xref="GI:42717"
/db_xref="GOA:P15905"
/translation="MTQLTPIQLFKMLSDETRLGIVLLRMGELCVCDLCLMALDQSQ PKTSRHLMRESGILLDRKQGWVHYRLSPHSWAAQIIIEQAWLSQQDDVQVIARK LASVNCSSSKAVCI"
482..511
/note="pot. stem-loop structure"

terminator
482..511
/note="pot. stem-loop structure"

ORIGIN
Query Match 100.0%; Score 28; DB 1; Length 727;
Best Local Similarity 100.0%; Pred. No. 0.79;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CAACACATACCAAAACGCATATGATT 28
Db 71 CAACACATACCAAAACGCATATGATT 44

RESULT 2
AP005147/c 120826 bp DNA circular BCT 14-MAY-2002
LOCUS Salmonella typhimurium plasmid R64 DNA, complete sequence.
DEFINITION AP005147
ACCESSION AP005147
VERSION AP005147.1 GI:20521502
KEYWORDS
SOURCE Salmonella typhimurium
ORGANISM Salmonella typhimurium
Bacteria: Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Salmonella.

REFERENCE 1
AUTHORS Komano, T., Kubo, A. and Nishioka, T.
TITLE Shufflon: multi-inversion of four contiguous DNA segments of plasmid R64 creates seven different open reading frames
JOURNAL Nucleic Acids Res. 15 (3), 1165-1172 (1987)
MEDLINE 87146423
PUBMED 3029698
REFERENCE 2
AUTHORS Kubo, A., Kusukawa, A. and Komano, T.
TITLE Nucleotide sequence of the rci gene encoding shufflon-specific DNA recombinase in the IncII plasmid R64: homology to the site-specific recombinases of integrase family

JOURNAL Mol. Gen. Genet. 213 (1), 30-35 (1988)
MEDLINE 89127142
PUBMED 3065610
REFERENCE 3
AUTHORS Komano, T., Toyoshima, A., Morita, K. and Nishioka, T.
TITLE Cloning and nucleotide sequence of the *oriT* region of the IncII plasmid R64
JOURNAL J. Bacteriol. 170 (9), 4385-4387 (1988)
MEDLINE 88314948
PUBMED 3045094
REFERENCE 4
AUTHORS Furuya, N., Nishioka, T. and Komano, T.
TITLE Nucleotide sequence and functions of the *oriT* operon in IncII plasmid R64
JOURNAL J. Bacteriol. 173 (7), 2231-2237 (1991)
MEDLINE 91177811
PUBMED 1848841
REFERENCE 5
AUTHORS Furuya, N. and Komano, T.
TITLE Determination of the nick site at *oriT* of IncII plasmid R64: global similarity of *oriT* structures of IncII and IncP plasmids
JOURNAL J. Bacteriol. 173 (20), 6612-6617 (1991)
MEDLINE 92011438
PUBMED 1917882
REFERENCE 6
AUTHORS Kim, S.R., Funayama, N. and Komano, T.
TITLE Nucleotide sequence and characterization of the *traABCD* region of IncII plasmid R64
JOURNAL J. Bacteriol. 175 (16), 5035-5042 (1993)
MEDLINE 93352408
PUBMED 8349545
REFERENCE 7
AUTHORS Furuya, N. and Komano, T.
TITLE Surface exclusion gene of IncII plasmid R64: nucleotide sequence and analysis of deletion mutants
JOURNAL Plasmid 32 (1), 80-84 (1994)
MEDLINE 95083745
PUBMED 7591676
REFERENCE 8
AUTHORS Furuya, N. and Komano, T.
TITLE Nucleotide sequence and characterization of the *trbABC* region of the IncII plasmid R64: existence of the *pnd* gene for plasmid maintenance within the transfer region
JOURNAL J. Bacteriol. 178 (6), 1491-1497 (1996)
MEDLINE 96198148
PUBMED 8626273
REFERENCE 9
AUTHORS Kim, S.R. and Komano, T.
TITLE The plasmid R64 thin pilus identified as a type IV pilus
JOURNAL J. Bacteriol. 179 (11), 3594-3603 (1997)
MEDLINE 97315231
PUBMED 9171405
REFERENCE 10
AUTHORS Narahara, K., Rahman, E., Furuya, N. and Komano, T.
TITLE Requirement of a limited segment of the *sog* gene for plasmid R64 conjugation
JOURNAL Plasmid 38 (1), 1-11 (1997)
MEDLINE 97428559
PUBMED 9281491
REFERENCE 11
AUTHORS Furuya, N. and Komano, T.
TITLE Mutational analysis of the R64 *oriT* region: requirement for precise location of the Nika-binding sequence
JOURNAL J. Bacteriol. 179 (23), 7291-7297 (1997)
MEDLINE 98053841
PUBMED 9393692
REFERENCE 12
AUTHORS Yoshida, T., Furuya, N., Ishikura, M., Isobe, T., Haino-Fukushima, K., Ogawa, T. and Komano, T.
TITLE Purification and characterization of thin pilus of IncII plasmid Colib-P9 and R64: formation of Pili-specific cell aggregates by type IV pilus
JOURNAL J. Bacteriol. 180 (11), 2842-2848 (1998)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 26, 2004, 15:20:44 ; Search time 584.231 Seconds
(without alignments)
3115.905 Million cell updates/sec

Title: US-10-676-299-5

Perfect score: 42

Sequence: 1 ctgcaactacatttgctta.....tcataatgttttgactta 42

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_htg.*

3: gb_in.*

4: gb_ov.*

5: gb_ro.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vl.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_or.*

21: em_ov.*

22: em_pat.*

23: em_ph.*

24: em_pl.*

25: em_ro.*

26: em_sts.*

27: em_un.*

28: em_vl.*

29: em_vl.*

30: em_htg_hum.*

31: em_htg_inv.*

32: em_htg_other.*

33: em_htg_mus.*

34: em_htg_pln.*

35: em_htg_rod.*

36: em_htg_mam.*

37: em_htg_vrt.*

38: em_sy.*

39: em_htgo_hum.*

40: em_htgo_mus.*

41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	42	100.0	3492	1	ECARSRC	X80057 E. coli gene
2	42	100.0	10240	1	AE000426	AE000426 Escherich
3	42	100.0	11524	1	AE015361	AE015361 Shigella
4	42	100.0	179941	2	AC145934	AC145934 Gallus ga
5	42	100.0	225419	1	ECOW76	U00039 E. coli chr
6	42	100.0	242495	2	AC146183	AC146183 Pan trogl
7	42	100.0	289816	1	AE016992	AE016992 Shigella
8	37.2	88.6	11071	1	AE005575	AE005575 Escherich
9	37.2	88.6	267888	1	AE016768	AE016768 Escherich
10	33	78.6	901660	1	AF002565	AF002565 Escherich
11	25.6	61.0	94650	10	AF242431S2	AF242432 Mus muscu
12	25.6	61.0	198831	2	AC116741	AC116741 Mus muscu
13	25	59.5	88013	9	AC090511	AC090511 Homo sapi
14	25	59.5	150965	9	AC091915	AC091915 Homo sapi
15	25	59.5	169377	2	AC110578	AC110578 Homo sapi
16	25	59.5	178127	2	AC079076	AC079076 Homo sapi
17	25	59.5	198719	2	AC015716	AC015716 Homo sapi
18	24.6	58.6	149928	2	AC134950	AC134950 Danio rer
19	24.6	58.6	166112	2	AC132260	AC132260 Mus muscu
20	24.6	58.6	176496	10	AL772194	AL772194 Mouse DNA
21	24.6	58.6	196888	2	BX323087	BX323087 Danio rer
22	24.4	58.1	183699	2	BX255894	BX255894 Danio rer
23	24.4	58.1	221969	2	AC116134	AC116134 Mus muscu
24	24.2	57.6	156569	2	AC110190	AC110190 Homo sapi
25	24.2	57.6	164314	9	AC091691	AC091691 Homo sapi
26	24.2	57.6	194366	5	AL954179	AL954179 Zebrafish
27	24	57.1	447	11	BV006936	BV006936 sa09all.y
28	24	57.1	70329	2	AC087718	AC087718 Homo sapi
29	24	57.1	109891	9	AL353897	AL353897 Human DNA
30	24	57.1	162761	2	AC079916	AC079916 Homo sapi
31	24	57.1	198913	2	AC044915	AC044915 Homo sapi
32	24	57.1	207471	10	AC122249	AC122249 Mus muscu
33	24	57.1	212691	2	AC118594	AC118594 Mus muscu
34	24	57.1	272717	2	AC098469	AC098469 Rattus no
35	23.8	56.7	214795	2	AC110449	AC110449 Rattus no
36	23.8	56.7	233330	2	AC112548	AC112548 Rattus no
37	23.6	56.2	87548	5	BX323038	BX323038 Zebrafish
38	23.6	56.2	101904	5	AL935310	AL935310 Zebrafish
39	23.6	56.2	122100	8	AP005774	AP005774 Oryza sat
40	23.6	56.2	135070	2	AP003741	AP003741 Oryza sat
41	23.6	56.2	142711	10	AC121863	AC121863 Mus muscu
42	23.6	56.2	152433	2	AP004297	AP004297 Oryza sat
43	23.6	56.2	164936	5	BX255954	BX255954 Zebrafish
44	23.6	56.2	168990	2	AC122516	AC122516 Mus muscu
45	23.6	56.2	177205	2	AC079217	AC079217 Mus muscu

ALIGNMENTS

RESULT 1	ECARSRC	E.coli genes arsr, arsb, arsc.	3492 bp	DNA	linear	BCT 20-JUL-1995
LOCUS	ECARSRC	E.coli genes arsr, arsb, arsc.				
DEFINITION	X80057					
ACCESSION	X80057.1	GI:510824				
VERSION	arsB gene; arsc gene; arsenate reductase; arsenic-efflux pump;					
KEYWORDS	arsenic-inducible repressor; arsr gene.					
SOURCE	Escherichia coli					
ORGANISM	Escherichia coli					
	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;					
	Enterobacteriaceae; Escherichia.					
REFERENCE	1					
AUTHORS	Diorio, C., Cai, J., Marmor, J., Shinder, R. and DuBow, M.S.					
TITLE	An Escherichia coli chromosomal ars operon homolog is functional in					

```

JOURNAL      arsenic detoxification and is conserved in gram-negative bacteria
MEDLINE      J. Bacteriol. 177 (8), 2050-2056 (1995)
PUBMED       95238276
REFERENCE    7721697
AUTHORS      2 (bases 1 to 3492)
TITLE        Diorio, C.
JOURNAL      Direct Submission
SUBMITTED    (06-JUL-1994) C. Diorio, McGill University, 3775
UNIVERSITY   University Street, Montreal, Quebec, H3A 2B4, CANADA
FEATURES     Location/Qualifiers
SOURCE       1..3492
             /organism="Escherichia coli"
             /mol_type="genomic DNA"
             /strain="K12"
             /db_xref="taxon:562"
             repeat_region 528..538
             /rpt_type=INVERTED
             repeat_region 544..554
             /rpt_type=INVERTED
             gene          555..1061
             /gene="arsR"
             -35_signal    555..660
             /gene="arsR"
             -10_signal    678..683
             /gene="arsR"
             RBS          697..701
             /gene="arsR"
             CDS          708..1061
             /gene="arsR"
             /codon_start=1
             /transl_table=1
             /product="arsenic-inducible repressor"
             /protein_id="CAA56361.1"
             /db_xref="GI:510825"
             /db_xref="GOA:P37309"
             /db_xref="SWISS-PROT:P37309"
             /translation="MSFLPIQLFKILADETRIGIVLLSELGELCVCDICTALDOSQ
             PKISHALLRESGLLDRKQKWHYELSHPIPAWAAKIIDEAWRCEQEKQVQIYRN
             LARQCSGDSKNICS"
             1071..2404
             /gene="arsB"
             -35_signal    1071..1076
             /gene="arsB"
             repeat_region 1072..1084
             /rpt_type=INVERTED
             repeat_region 1087..1099
             /rpt_type=INVERTED
             -10_signal    1088..1093
             /gene="arsB"
             RBS          1106..1110
             /gene="arsB"
             CDS          1115..2404
             /gene="arsB"
             /codon_start=1
             /transl_table=1
             /product="membrane-located arsenite efflux pump"
             /protein_id="CAA56362.1"
             /db_xref="GI:516211"
             /db_xref="GOA:P37310"
             /db_xref="SWISS-PROT:P37310"
             /translation="MLLAGAIFVITIVLWQPKGLGWSATLGAVALVGVVHPG
             DIPVYNTVMNATAFIATVILIIISLLDSGFEWALHVRWNGRGLLFTWVLLG
             RAAALFANDGAATLPIVIAMLALGFSKGTTLAFVMAAGFIADTASLPIVSNLV
             NIVSADFGLGFREYASVMVVDIAATVATVLMHLVFRKDIPONYDMALLKSPALV
             KDPATFKTGVVLLLVGVFFVPLGPIPVSAIAVAGVALIFVVAKRGHAINTKVLR
             GAPWQIVFSLGMLVAVYGLNAGLTVLSGLNVLADNGLMAATLGTGPTLAFLSI
             MNMTVLAVGLSDGSTASGVIKEMVYANVIGCDLGPKITPIGSLATLILWLHLVLSQ
             KNTISWYVYRTGINTLPVLFVTLAALALRSLFTL"
             2417..2842
             /gene="arsC"
             2417..2842
             /gene="arsC"
             /codon_start=1

```

```

/transl_table=1
/product="arsenate reductase"
/protein_id="CAA56363.1"
/db_xref="GI:516212"
/db_xref="GOA:P37311"
/db_xref="SWISS-PROT:P37311"
/translation="MSNITVYHNPACGTSRNTLEMRNSGTEPTIIHYLETPTPTDEL
VKLIADMGISVRALLKKNVPEELGLAEDKTDRLIDFMLQHPILINRPVVTPLG
TLCRSPSEVLEILPDAQGAFSGKEDGKXVDEAGKELK"
repeat_region 2892..2905
/rpt_type=INVERTED
repeat_region 2910..2923
/rpt_type=INVERTED
ORIGIN
Query Match      100.0%; Score 42; DB 1; Length 3492;
Best Local Similarity 100.0%; Pred. No. 0.00044;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CTGCACCTACACATTCGTTAAGTCATATATATGTTTGGACTTA 42
DB 621 CTGCACCTACACATTCGTTAAGTCATATATGTTTGGACTTA 662
RESULT 2
AE000426      10240 bp      DNA      linear      BCT 01-DEC-2000
LOCUS        Escherichia coli K12 MG1655 section 316 of 400 of the complete
DEFINITION   genome.
ACCESSION    AE000426 U00096
VERSION      AE000426.1 GI:1789910
KEYWORDS     Escherichia coli K12
SOURCE       Escherichia coli K12
ORGANISM     Escherichia coli K12
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.
REFERENCE    1 (Bases 1 to 10240)
AUTHORS      Blattner, F.R., Plunkett, G. III, Bloch, C.A., Perna, N.T., Burland, V.,
Riley, M., Collado-Vides, J., Glasner, J.D., Rode, C.K., Mayhew, G.F.,
Gregor, J., Davis, N.W., Kirkpatrick, H.A., Goeden, M.A., Rose, D.J.,
Mau, B. and Shao, Y.
The complete genome sequence of Escherichia coli K-12
Science 277 (5331), 1453-1474 (1997)
97426617
PUBMED       9278503
REFERENCE    2 (Bases 1 to 10240)
AUTHORS      Blattner, F.R.
Direct Submission
Submitted (16-JAN-1997) Guy Plunkett III, Laboratory of Genetics,
University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA.
Email: ecollegenetics.wisc.edu Phone: 608-262-2534 Fax:
608-263-7459
3 (Bases 1 to 10240)
AUTHORS      Blattner, F.R.
Direct Submission
Submitted (02-SEP-1997) Guy Plunkett III, Laboratory of Genetics,
University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA.
Email: ecollegenetics.wisc.edu Phone: 608-262-2534 Fax:
608-263-7459
4 (Bases 1 to 10240)
AUTHORS      Plunkett, G. III.
Direct Submission
Submitted (13-OCT-1998) Laboratory of Genetics, University of
Wisconsin, 445 Henry Mall, Madison, WI 53706, USA
This sequence was determined by the E. coli Genome Project at the
University of Wisconsin-Madison (Frederick R. Blattner, director).
Supported by NIH grants HG00301 and HG01428 (from the Human Genome
Project and NCRR). The entire sequence was independently
determined from E. coli K12 strain MG1655. Predicted open reading
frames were determined using GeneMark software, kindly supplied by
Mark Borodovsky, Georgia Institute of Technology, Atlanta, GA,
30332 [e-mail: mark@amber.gatech.edu]. Open reading frames that
have been correlated with genetic loci are being annotated with CG

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 26, 2004, 15:20:44 ; Search time 556.41 Seconds
(without alignments)
3115.905 Million cell updates/sec

Title: US-10-676-299-6
Perfect score: 40
Sequence: 1 taagtcaaaacatatatgacttaacgaatgtgaagtgc 40

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*

- 1: gb_pa.*
- 2: gb_hgt.*
- 3: gb_in.*
- 4: gb_em.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_or.*
- 21: em_ov.*
- 22: em_ph.*
- 23: em_pat.*
- 24: em_pl.*
- 25: em_ro.*
- 26: em_sts.*
- 27: em_un.*
- 28: em_vi.*
- 29: em_hgt_hum.*
- 30: em_hgt_inv.*
- 31: em_hgt_inv.*
- 32: em_hgt_inv.*
- 33: em_hgt_mus.*
- 34: em_hgt_pln.*
- 35: em_hgt_rod.*
- 36: em_hgt_mam.*
- 37: em_hgt_vrt.*
- 38: em_sy.*
- 39: em_hgt_hum.*
- 40: em_hgt_mus.*
- 41: em_hgtg_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
C 1	40	100.0	3492	1	ECARSRC	X80057 E.coli gene
C 2	40	100.0	10240	1	AE000426	AE000426 Escherich
C 3	40	100.0	11524	1	AE015361	AE015361 Shigella
C 4	40	100.0	179941	2	AC145934	AC145934 Gallus ga
C 5	40	100.0	225419	1	ECOUW76	UC00039 E. coli chr
C 6	40	100.0	242495	2	AC146183	AC146183 Pan trogl
C 7	40	100.0	289816	1	AE016992	AE016992 Shigella
C 8	35.2	88.0	11071	1	AE005575	AE005575 Escherich
C 9	35.2	88.0	267888	1	AF002565	AF002565 Escherich
C 10	31.6	79.0	301660	1	AE016768	AE016768 Escherich
C 11	24.6	61.5	161112	2	AC102260	AC102260 Mus muscu
C 12	24.6	61.5	176436	10	AL772194	AL772194 Mouse DNA
C 13	24.2	60.5	90650	10	AF242432	AF242432 Mus muscu
C 14	24.2	60.5	158569	2	AC110190	AC110190 Homo sapi
C 15	24.2	60.5	164314	9	AC091691	AC091691 Homo sapi
C 16	24.2	60.5	194366	5	AF954179	AF954179 Zebrafish
C 17	24.2	60.5	198631	2	AC116741	AC116741 Mus muscu
C 18	24	60.0	88013	9	AC094511	AC094511 Homo sapi
C 19	24	60.0	150965	9	AC091515	AC091515 Homo sapi
C 20	24	60.0	169377	2	AC110570	AC110570 Homo sapi
C 21	24	60.0	178127	2	AC079076	AC079076 Homo sapi
C 22	24	60.0	198719	2	AC015716	AC015716 Homo sapi
C 23	24	60.0	207471	10	AC122249	AC122249 Mus muscu
C 24	24	60.0	213691	2	AC118594	AC118594 Mus muscu
C 25	23.8	59.5	214795	2	AC110449	AC110449 Rattus no
C 26	23.8	59.5	233330	2	AC112548	AC112548 Rattus no
C 27	23.6	59.0	87548	5	BX323038	BX323038 Zebrafish
C 28	23.6	59.0	101904	5	AL935310	AL935310 Zebrafish
C 29	23.6	59.0	122100	8	AF005774	AF005774 Oryza sat
C 30	23.6	59.0	135070	2	AP003741	AP003741 Oryza sat
C 31	23.6	59.0	142711	10	AC121863	AC121863 Mus muscu
C 32	23.6	59.0	149928	2	AC134950	AC134950 Danio rer
C 33	23.6	59.0	152433	2	AP004297	AP004297 Oryza sat
C 34	23.6	59.0	164936	5	BX255954	BX255954 Zebrafish
C 35	23.6	59.0	168990	2	AC122516	AC122516 Mus muscu
C 36	23.6	59.0	183417	5	AL845320	AL845320 Zebrafish
C 37	23.6	59.0	184808	2	AC118619	AC118619 Mus muscu
C 38	23.6	59.0	186668	2	AC114627	AC114627 Mus muscu
C 39	23.6	59.0	196688	2	BX323087	BX323087 Danio rer
C 40	23.6	59.0	209383	2	AC133953	AC133953 Mus muscu
C 41	23.6	59.0	221146	10	AC098739	AC098739 Mus muscu
C 42	23.6	59.0	224806	2	BX004858	BX004858 Danio rer
C 43	23.6	59.0	225082	10	BX548065	BX548065 Mouse DNA
C 44	23.6	59.0	226871	2	AC098604	AC098604 Rattus no
C 45	23.6	59.0	253066	2	AL935184	AL935184 Danio rer

ALIGNMENTS

RESULT 1	ECARSRC/c	3492 bp	DNA	linear	BCT 20-JUL-1995
LOCUS	ECARSRC				
DEFINITION	E.coli genes arsr, arsb, arsc.				
ACCESSION	X80057.1	GI:510824			
VERSION	arsB gene; arsc gene; arsenate reductase; arsenic-efflux pump;				
KEYWORDS	arsenic-inducible repressor; arsr gene.				
SOURCE	Escherichia coli				
ORGANISM	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Escherichia.				
REFERENCE	1				
AUTHORS	Diorio, C., Cai, J., Marmor, J., Shinder, R. and DuBow, M.S.				
TITLE	An Escherichia coli chromosomal ars operon homolog is functional in				

arsenic detoxification and is conserved in gram-negative bacteria
J. Bacteriol. 177 (8), 2050-2056 (1995)

```
JOURNAL
MEDLINE
PUBMED
REFERENCE 2 (bases 1 to 3492)
AUTHORS
Diorio, C.
TITLE
Submitted (06-JUL-1994) C. Diorio, McGill University, 3775
JOURNAL
University Street, Montreal, Quebec, H3A 2B4, CANADA
FEATURES
source
1. 3492
/organism="Escherichia coli"
/mol_type="genomic DNA"
/strain="K12"
/db_xref="taxon:562"
528..538
repeat_region
544..554
repeat_region
555..1061
/rpt_type=INVERTED
655..660
gene
655..660
-35_signal
678..683
-10_signal
697..701
RBS
708..1061
CDS
/codon_start=1
/transl_table=11
/product="arsenic-inducible repressor"
/protein_id="CAA56362.1"
/db_xref="GI:510825"
/db_xref="GOA:P37309"
/transl_table="MSFLPIQLFKILADETRHGLVLLSELGVCVDLCTALDQSQ
PKSRRLALIREGLIDRRQKQWVHRLSPHAWAAKIIDEARWCEQKQVQALVRN
LARQNGSDSKNICS"
1071..2404
/gene="arsB"
1071..1076
-35_signal
1072..1084
repeat_region
1087..1099
repeat_region
1088..1093
-10_signal
1106..1110
RBS
1115..2404
CDS
/codon_start=1
/transl_table=11
/product="membrane-located arsenite efflux pump"
/protein_id="CAA56362.1"
/db_xref="GI:516211"
/db_xref="GOA:P37310"
/transl_table="MLLAGAIFVLTIVLWOPKGLIGWSATLGAVALVGVVHPG
DIPVNVNATAATIAVIIISLLDESGFTEWALHVSRWNGRGLLFTWVLLG
AAVAALFANPAGAAIITPIVIMLLALGSKGTTIAFVAAGFIADTASLPLTNSLV
NIVSADPEGLGFEYASVMVPDIAIATVAVLHLYFKDIPQNYDVALLKSPAEAI
KDPAFTGHWVLLLLVGFVLEPLGIPVSAIAGVALLIFVAKRGHAINTKVLR
GAPQIVIFSGITLVVYGLRNAGLTYSGLVNLVADNLWAAATGCTGFTAFVSSI
MNNPTVLVGLSIDSTASGVKEAMVYANVIGCDLGPKITPIGSLATLLWLHLVLSQ
KNMTISGYYFRGTGIINTLPVLVFTVLAALRLSFTL"
2417..2842
/gene="arsC"
2417..2842
CDS
/gene="arsC"
/codon_start=1
```

```
/transl_table=11
/product="arsenate reductase"
/protein_id="CAA56363.1"
/db_xref="GI:516212"
/db_xref="GOA:P37311"
/db_xref="SWISS-PROT:P37311"
/transl_table="MSNITIVNPACGTSNTLEMRNSGTEPTIIHYLETPTPRDEL
VKLIADMGISVRLAKNKVEYVEELGAEKFTDRLIDMLQHPILNRPVVTPLG
TELCPSEVLEILPDAQKGAFSKEDGKVVDEAGKELK"
2892..2905
repeat_region
2910..2923
/rpt_type=INVERTED
2910..2923
repeat_region
2910..2923
/rpt_type=INVERTED
ORIGIN
Query Match 100.0%; Score 40; DB 1; Length 3492;
Best Local Similarity 100.0%; Pred. No. 0.0022;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TAAGTCAAAACATATATGACTTAACGAATGTGAAGTGC 40
|||||
Db 662 TAAGTCAAAACATATATGACTTAACGAATGTGAAGTGC 623
RESULT 2
AE000426/c
LOCUS
DEFINITION
Escherichia coli K12 MGI655 section 316 of 400 of the complete
genome.
ACCESSION
AE000426 U00096
VERSION
AE000426.1 GI:1789910
KEYWORDS
Escherichia coli K12
SOURCE
Escherichia coli K12
ORGANISM
Escherichia coli K12
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.
REFERENCE
1 (bases 1 to 10240)
Blattner, F.R., Plunkett, G. III, Bloch, C.A., Perna, N.T., Burland, V.,
Riley, M., Collado-Vides, J., Glasner, J.D., Rode, C.K., Mayhew, G.F.,
Gregor, J., Davis, N.W., Kirkpatrick, H.A., Goeden, M.A., Rose, D.J.,
Mau, B. and Shao, Y.
The complete genome sequence of Escherichia coli K-12
Science 277 (8331), 1453-1474 (1997)
97426617
PUBMED
9278503
REFERENCE 2 (bases 1 to 10240)
Blattner, F.R.
Direct Submission
Submitted (16-JAN-1997) Guy Plunkett III, Laboratory of Genetics,
University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA.
Email: ecolig@genetics.wisc.edu Phone: 608-262-2534 Fax:
608-263-7459
3 (bases 1 to 10240)
Blattner, F.R.
Direct Submission
Submitted (02-SEP-1997) Guy Plunkett III, Laboratory of Genetics,
University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA.
Email: ecolig@genetics.wisc.edu Phone: 608-262-2534 Fax:
608-263-7459
4 (bases 1 to 10240)
Blattner, F.R.
Direct Submission
Submitted (13-OCT-1998) Laboratory of Genetics, University of
Wisconsin, 445 Henry Mall, Madison, WI 53706, USA
This sequence was determined by the E. coli Genome Project at the
University of Wisconsin-Madison (Frederick R. Blattner, director).
Supported by NIH grants HG00301 and HG01428 (from the Human Genome
Project and NCHGR). The entire sequence was independently
determined from E. coli K12 strain MGI655. Predicted open reading
frames were determined using GeneMark software, kindly supplied by
Mark Borodovsky, Georgia Institute of Technology, Atlanta, GA.
30332 [e-mail: mark@amber.gatech.edu]. Open reading frames that
have been correlated with genetic loci are being annotated with CG
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 26, 2004, 15:20:44 ; Search time 319.936 Seconds
(without alignments)
3115.905 Million cell updates/sec

Title: US-10-676-299-7

Perfect score: 23

Sequence: 1 ttaatcatatgggttttggta 23

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: . 3470272 segs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.cm.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sy.*

13: gb.un.*

14: gb.vi.*

15: em.ba.*

16: em.fun.*

17: em.hum.*

18: em.in.*

19: em.mu.*

20: em.om.*

21: em.or.*

22: em.ov.*

23: em.pat.*

24: em.ph.*

25: em.pl.*

26: em.ro.*

27: em.sts.*

28: em.un.*

29: em.vi.*

30: em.htg.hum.*

31: em.htg.inv.*

32: em.htg.other.*

33: em.htg.mus.*

34: em.htg.pln.*

35: em.htg.rod.*

36: em.htg.mam.*

37: em.htg.vrt.*

38: em.sy.*

39: em.htgo.hum.*

40: em.htgo.mus.*

41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	23	100.0	727	1	ECRFARS	X16045 E. coli R-f
2	23	100.0	120826	1	AP005147	AP005147 Salmonella
3	19	82.6	176325	2	BX279525	BX279525 Danio rer
4	19	82.6	184543	5	BX571681	BX571681 Zebrafish
5	19	82.6	215891	5	AL928692	AL928692 Zebrafish
6	19	82.6	219601	2	BX649641	BX649641 Danio rer
7	19	82.6	253172	5	BX571883	BX571883 Zebrafish
8	18.8	81.7	620	3	AF129331	AF129331 Phrantela
9	18.8	81.7	3727	6	AX713938	AX713938 Sequence
10	18.8	81.7	3727	9	AK055877	AK055877 Homo sapi
11	18.8	81.7	41447	1	AE011312	AE011312 Leptospir
12	18.8	81.7	38574	3	AF314193	AF314193 Drosophil
13	18.8	81.7	74881	2	AC020320	AC020320 Drosophil
14	18.8	81.7	112659	9	AC010677	AC010677 Homo sapi
15	18.8	81.7	119595	5	BX005186	BX005186 Zebrafish
16	18.8	81.7	137635	2	BX629344	BX629344 Danio rer
17	18.8	81.7	161411	2	AC102700	AC102700 Mus muscu
18	18.8	81.7	170869	3	AC011696	AC011696 Drosophil
19	18.8	81.7	171831	3	AC007473	AC007473 Drosophil
20	18.8	81.7	193708	9	AC005035	AC005035 Homo sapi
21	18.8	81.7	223489	5	BX248180	BX248180 Zebrafish
22	18.8	81.7	228048	2	AC122883	AC122883 Mus muscu
23	18.8	81.7	261690	3	AE003825	AE003825 Drosophil
24	18.8	81.7	269223	6	AF408762	AF408762 Sequence
25	18.8	81.7	269223	6	AX067466	AX067466 Sequence
26	18.8	81.7	295094	2	BX649429	BX649429 Danio rer
27	18.4	80.0	21479	3	CEFG03E09	Z79756 Caenorhabd
28	18.4	80.0	34796	3	CEF53C11	AX224188 Medicago
29	18.4	80.0	99431	8	AY224188	AC114585 Medicago
30	18.4	80.0	124457	8	AC145585	AC111549 Rattus no
31	18.4	80.0	212669	2	AC111549	AC094788 Rattus no
32	18.4	80.0	261226	2	AC094788	AJ598878 Arabidops
33	18.2	79.1	201	8	AJ598878	AY201116 Arabidops
34	18.2	79.1	430	8	AY201116	BX530054 Arabidops
35	18.2	79.1	460	11	BX530054	AJ506884 Buthus oc
36	18.2	79.1	466	3	BOC506884	AJ506893 Buthus oc
37	18.2	79.1	466	3	BOC506893	AJ506894 Buthus oc
38	18.2	79.1	466	3	BOC506894	AJ506897 Buthus oc
39	18.2	79.1	466	3	BOC506896	AJ506897 Buthus oc
40	18.2	79.1	466	3	BOC506897	AJ506901 Buthus oc
41	18.2	79.1	466	3	BOC506901	AJ506902 Buthus oc
42	18.2	79.1	466	3	BOC506902	BX530044 Arabidops
43	18.2	79.1	486	11	BX530044	AJ550701 Mesobuthu
44	18.2	79.1	491	3	MEU550701	AJ550702 Mesobuthu
45	18.2	79.1	491	3	MEU550702	

ALIGNMENTS

RESULT 1
ECRFARS ECRFARS E. coli R-factor R773 arsr gene.
LOCUS X16045
DEFINITION X16045.1 GI:42716
ACCESSION
VERSION
KEYWORDS arsenical resistance; arsr gene; Arsr protein; DNA-binding protein; regulatory protein; resistance gene.
SOURCE Escherichia coli
ORGANISM Escherichia coli
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Escherichia.
REFERENCE 1 (bases 1 to 727)
AUTHORS San Francisco,M.J., Hope,C.L., Owolabi,J.B., Tisa,L.S. and Rosen,B.P.

TITLE Identification of the metalloregulatory element of the plasmid-encoded arsenical resistance operon

JOURNAL MEDLINE Nucleic Acids Res. 18 (3), 619-624 (1990)

PUBMED 90174986

REFERENCE 2 2408017

AUTHORS Rosen, B.P.

TITLE Direct Submission

JOURNAL Submitted (09-AUG-1989) Rosen B.P., Department of Biochemistry, Wayne State University, School of Medicine, 54- E Canfield Avenue, Detroit MI 48201, U S A

FEATURES

source

1. 727

/organism="Escherichia coli"

/mol_type="genomic DNA"

/db_xref="taxon:562"

/clone="pMSU1"

/clone_lib="pSR322"

73. 79

/note="pot. -35 region"

96. 102

/note="pot. -10 region"

107

promoter

promoter

misc_feature

114. 118

/note="transcriptional start site"

125. 478

/note="pot. ribosome binding site"

/note="unnamed protein product; ArsR protein (AA 1 - 117)"

RBS

CDS

/codon_start=1

/transl_table=11

/protein_id="CAA34168.1"

/db_xref="GI:42717"

/db_xref="GOA:P15905"

/db_xref="SWISS-PROT:P15905"

/translation="MLQTLPLQFLKNSDSTRIGIVLLREMGELVCVDCILCMALDOSQ PKISPHLAMERESGILLDRKQGWVHRLSPHPSWAQILQEWLSSQDDVQVIAK LASVNCSSGSKAVCI"

482. 511

terminator

/note="pot. stem-loop structure"

ORIGIN

Query Match 100.0%; Score 23; DB 1; Length 727;

Best Local Similarity 100.0%; Pred. No. 26;

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTAATCATATCGTTTGGTTA 23

Db 42 TTAATCATATCGTTTGGTTA 64

RESULT 2

AP005147

LOCUS AP005147 120826 bp DNA circular 3014-MAY-2002

DEFINITION *Salmonella typhimurium* plasmid R64 DNA, complete sequence.

ACCESSION AP005147

VERSION AP005147.1 GI:20521502

KEYWORDS *Salmonella typhimurium*

SOURCE *Salmonella typhimurium*

ORGANISM *Salmonella typhimurium*

REFERENCE 1

AUTHORS Komano, T., Kubo, A., and Nishioka, T.

TITLE Shufflon: multi-inversion of four contiguous DNA segments of plasmid R64 creates seven different open reading frames

JOURNAL Nucleic Acids Res. 15 (3), 1165-1172 (1987)

MEDLINE 87146423

PUBMED 3029698

REFERENCE 2

AUTHORS Kubo, A., Kusukawa, A., and Komano, T.

TITLE Nucleotide sequence of the rci gene encoding shufflon-specific DNA recombinase in the IncII plasmid R64: homology to the site-specific recombinases of integrase family

JOURNAL MEDLINE Mol. Gen. Genet. 213 (1), 30-35 (1988)

PUBMED 89127142

REFERENCE 3 3065610

AUTHORS Komano, T., Toyoshima, A., Morita, K., and Nishioka, T.

TITLE Cloning and nucleotide sequence of the *oriT* region of the IncII plasmid R64

JOURNAL J. Bacteriol. 170 (9), 4385-4387 (1988)

MEDLINE 88314948

PUBMED 3045094

REFERENCE 4

AUTHORS Furuya, N., Nishioka, T., and Komano, T.

TITLE Nucleotide sequence and functions of the *oriT* operon in IncII plasmid R64

JOURNAL J. Bacteriol. 173 (7), 2231-2237 (1991)

MEDLINE 91177811

PUBMED 1848841

REFERENCE 5

AUTHORS Furuya, N., and Komano, T.

TITLE Determination of the nick site at *oriT* of IncII plasmid R64: global similarity of *oriT* structures of IncII and IncP plasmids

JOURNAL J. Bacteriol. 173 (20), 6612-6617 (1991)

MEDLINE 92011438

PUBMED 1917882

REFERENCE 6

AUTHORS Kim, S.R., Funayama, N., and Komano, T.

TITLE Nucleotide sequence and characterization of the *traABCD* region of IncII plasmid R64

JOURNAL J. Bacteriol. 175 (16), 5035-5042 (1993)

MEDLINE 93352408

PUBMED 8349545

REFERENCE 7

AUTHORS Furuya, N., and Komano, T.

TITLE Surface exclusion gene of IncII plasmid R64: nucleotide sequence and analysis of deletion mutants

JOURNAL Plasmid 32 (1), 80-84 (1994)

MEDLINE 95083745

PUBMED 7991676

REFERENCE 8

AUTHORS Furuya, N., and Komano, T.

TITLE Nucleotide sequence and characterization of the *trbABC* region of the IncII plasmid R64: existence of the *pnd* gene for plasmid maintenance within the transfer region

JOURNAL J. Bacteriol. 178 (6), 1491-1497 (1996)

MEDLINE 96198148

PUBMED 8626273

REFERENCE 9

AUTHORS Kim, S.R., and Komano, T.

TITLE The plasmid R64 thin pilus identified as a type IV pilus

JOURNAL J. Bacteriol. 179 (11), 3594-3603 (1997)

MEDLINE 97315231

PUBMED 9171405

REFERENCE 10

AUTHORS Narahara, K., Rahman, E., Furuya, N., and Komano, T.

TITLE Requirement of a limited segment of the *sog* gene for plasmid R64 conjugation

JOURNAL Plasmid 38 (1), 1-11 (1997)

MEDLINE 97428559

PUBMED 9281491

REFERENCE 11

AUTHORS Furuya, N., and Komano, T.

TITLE Mutational analysis of the R64 *oriT* region: requirement for precise location of the *Nika*-binding sequence

JOURNAL J. Bacteriol. 179 (23), 7291-7297 (1997)

MEDLINE 98053841

PUBMED 9393692

REFERENCE 12

AUTHORS Yoshida, T., Furuya, N., Ishikura, M., Isobe, T., Haino-Fukushima, K., Ogawa, T., and Komano, T.

TITLE Purification and characterization of thin pili of IncII plasmids Colib-99 and R64: formation of PiliV-specific cell aggregates by type IV pili

JOURNAL J. Bacteriol. 180 (11), 2842-2848 (1998)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 26, 2004, 15:20:44 ; Search time 292.115 Seconds
(without alignments)
3115.905 Million cell updates/sec

Title: US-10-676-299-8
Perfect score: 21
Sequence: 1 taacacaaacgcataatgatt 21

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_scs.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vl.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_scs.*
- 28: em_un.*
- 29: em_vl.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_rdt.*
- 36: em_htg_mam.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_hngo_hum.*
- 40: em_hngo_mus.*
- 41: em_hngo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	21	100.0	727	1	ECRFARSR	X16045 E. coli R-f
C 2	21	100.0	120826	1	AP005147	AP005147 Salmonella
C 3	18.4	87.6	21479	3	CBRG03E09	AC084483 Caenorhabd
C 4	18.4	87.6	34796	3	CEP53C11	Z79756 Caenorhabd
C 5	18.4	87.6	39574	3	AF314193	AF314193 Drosophila
C 6	18.4	87.6	74881	2	AC020320	AC020320 Drosophila
C 7	18.4	87.6	170869	3	AC011696	AC011696 Drosophila
C 8	18.4	87.6	171831	3	AC007473	AC007473 Drosophila
C 9	18.4	87.6	261690	3	AE003825	AE003825 Drosophila
C 10	18	85.7	96444	9	HSJ738116	AL049867 Human DNA
C 11	18	85.7	19082	2	AC008057	AC008057 Homo sapi
C 12	18	85.7	163494	9	AC007630	AC007630 Homo sapi
C 13	17.8	84.8	466	3	BOC506884	AJ506884 Butus oc
C 14	17.8	84.8	466	3	BOC506893	AJ506893 Butus oc
C 15	17.8	84.8	466	3	BOC506894	AJ506894 Butus oc
C 16	17.8	84.8	466	3	BOC506896	AJ506896 Butus oc
C 17	17.8	84.8	466	3	BOC506897	AJ506897 Butus oc
C 18	17.8	84.8	466	3	BOC506901	AJ506901 Butus oc
C 19	17.8	84.8	466	3	BOC506902	AJ506902 Butus oc
C 20	17.8	84.8	491	3	MEU550701	AJ550701 Mesobuthu
C 21	17.8	84.8	491	3	MEU550702	AJ550702 Mesobuthu
C 22	17.8	84.8	491	3	MEU550703	AJ550703 Mesobuthu
C 23	17.8	84.8	599	3	AY128138	AY128138 Halgerda
C 24	17.8	84.8	599	3	AY128139	AY128139 Halgerda
C 25	17.8	84.8	603	3	MTAJ3267	AJ223267 Jorunna t
C 26	17.8	84.8	658	3	AF370829	AF370829 Androcton
C 27	17.8	84.8	676	3	AY116607	AY116607 Styela cl
C 28	17.8	84.8	778	5	AY169861	AY169861 Morella v
C 29	17.8	84.8	1263	3	S54796	S54796 cytochrome
C 30	17.8	84.8	10274	1	U32797	U32797 Haemophilus
C 31	17.8	84.8	14771	3	AB024528	AB024528 Halocynth
C 32	17.8	84.8	110000	6	BX005451_0	BX005451 Mus muscu
C 33	17.8	84.8	110000	6	AR274513_12	Continuation (13 o
C 34	17.8	84.8	127887	2	AP003688	AP003688 Oryza sat
C 35	17.8	84.8	127887	2	AC125477	AC125477 Medicago
C 36	17.8	84.8	131402	10	AL773583	AL773583 Mouse DNA
C 37	17.8	84.8	137635	2	EX629344	EX629344 Dario rer
C 38	17.8	84.8	162297	8	AP004317	AP004317 Oryza sat
C 39	17.8	84.8	165679	8	AC125475	AC125475 Medicago
C 40	17.8	84.8	177140	2	AP005612	AP005612 Oryza sat
C 41	17.8	84.8	233016	2	AC113719	AC113719 Rattus no
C 42	17.8	84.8	254644	2	AC136836	AC136836 Rattus no
C 43	17.8	84.8	256511	2	AC135040	AC135040 Rattus no
C 44	17.8	84.8	262721	2	AC106265	AC106265 Rattus no
C 45	17.8	84.8	270121	2	AC132060	AC132060 Rattus no

ALIGNMENTS

RESULT 1	ECRFARSR	727 bp	DNA	linear	BCT 07-SBP-1994
LOCUS	E. coli R-factor R773 arsr gene.				
DEFINITION	X16045				
ACCESSION	X16045.1	GI:42716			
VERSION	arsenical resistance; arsr gene; Arsr protein; DNA-binding protein;				
KEYWORDS	regulatory protein; resistance gene.				
SOURCE	Escherichia coli				
ORGANISM	Escherichia coli				
	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;				
	Enterobacteriaceae; Escherichia.				
REFERENCE	1 (bases 1 to 727)				
AUTHORS	San Francisco,M.J., Hope,C.L., Owolabi,J.B., Tisa,L.S. and				
	Rosen,B.P.				

TITLE Identification of the metalloregulatory element of the plasmid-encoded arsenical resistance operon

JOURNAL Nucleic Acids Res. 18 (3), 619-624 (1990)

MEDLINE 90174986

PUBMED 2408017

REFERENCE 2 (bases 1 to 727)

AUTHORS Rosen, B.P.

TITLE Direct Submission

JOURNAL Submitted (09-AUG-1989) Rosen B.P., Department of Biochemistry, Wayne State University, School of Medicine, 54- E Canfield Avenue, Detroit MI 48201, U S A

FEATURES

source 1..727 Location/Qualifiers

/organism="Escherichia coli"

/mol_type="genomic DNA"

/db_xref="taxon:562"

/clone="pWSU1"

/clone_lib="pBR322"

73..79

promoter /note="pot. -35 region"

96..102

promoter /note="pot. -10 region"

107

misc_feature /note="transcriptional start site"

114..118

RBS /note="pot. ribosome binding site"

125..478

CDS /note="unnamed protein product; ArsR protein (AA 1 - 117)"; /codon_start=1

/transl_table=1

/protein_id="CAA34168.1"

/db_xref="GI:42717"

/db_xref="GOA:P15905"

/db_xref="SWISS-PROT:P15905"

/translation="MLQTPQLPKNSDTRIGVLLRMGELCVCDLMDLDSQ PIRSHLAWRESGILLDRKQKWHVLSHPISWAAQIEQWLSCQDDVQVIARK LASVNCSSKAVCI"

482..511

terminator /note="pot. stem-loop structure"

ORIGIN

Query Match 100.0%; Score 21; DB 1; Length 727;

Best Local Similarity 100.0%; Pred. No. 58;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAACCAAAACGCATATGATT 21

Db 64 TAACCAAAACGCATATGATT 44

RESULT 2

AP005147/c 120826 bp DNA circular BCT 14-MAY-2002

LOCUS Salmonella typhimurium plasmid R64 DNA, complete sequence.

DEFINITION AP005147

ACCESSION AP005147.1

VERSION AP005147.1

KEYWORDS GI:20521502

SOURCE Salmonella typhimurium

ORGANISM Salmonella typhimurium

Bacteria: Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Salmonella.

REFERENCE 1

AUTHORS Komano, T., Kubo, A. and Nishioka, T.

TITLE Shufflon: multi-inversion of four contiguous DNA segments of plasmid R64 creates seven different open reading frames

JOURNAL Nucleic Acids Res. 15 (3), 1186-1172 (1987)

MEDLINE 87146423

PUBMED 3029698

REFERENCE 2

AUTHORS Kubo, A., Kusukawa, A. and Komano, T.

TITLE Nucleotide sequence of the rci gene encoding shufflon-specific DNA recombinase in the IncII plasmid R64: homology to the site-specific recombinases of integrase family

JOURNAL Mol. Gen. Genet. 213 (1), 30-35 (1988)

MEDLINE 89127142

PUBMED 3045610

REFERENCE 3

AUTHORS Komano, T., Toyoshima, A., Morita, K. and Nishioka, T.

TITLE Cloning and nucleotide sequence of the orit region of the IncII plasmid R64

JOURNAL J. Bacteriol. 170 (9), 4385-4387 (1988)

MEDLINE 88314948

PUBMED 3045094

REFERENCE 4

AUTHORS Furuya, N., Nishioka, T. and Komano, T.

TITLE Nucleotide sequence and functions of the orit operon in IncII plasmid R64

JOURNAL J. Bacteriol. 173 (7), 2231-2237 (1991)

MEDLINE 91177841

PUBMED 1848841

REFERENCE 5

AUTHORS Furuya, N. and Komano, T.

TITLE Determination of the nick site at orit of IncII plasmid R64: global similarity of orit structures of IncII and IncP plasmids

JOURNAL J. Bacteriol. 173 (20), 6612-6617 (1991)

MEDLINE 92011438

PUBMED 1917882

REFERENCE 6

AUTHORS Kim, S.R., Funayama, N. and Komano, T.

TITLE Nucleotide sequence and characterization of the traABCD region of IncII plasmid R64

JOURNAL J. Bacteriol. 175 (16), 5035-5042 (1993)

MEDLINE 93352408

PUBMED 8349545

REFERENCE 7

AUTHORS Furuya, N. and Komano, T.

TITLE Surface exclusion gene of IncII plasmid R64: nucleotide sequence and analysis of deletion mutants

JOURNAL Plasmid 32 (1), 80-84 (1994)

MEDLINE 95083745

PUBMED 7991676

REFERENCE 8

AUTHORS Furuya, N. and Komano, T.

TITLE Nucleotide sequence and characterization of the trbABC region of the IncII plasmid R64: existence of the end gene for plasmid maintenance within the transfer region

JOURNAL J. Bacteriol. 178 (6), 1491-1497 (1996)

MEDLINE 96198148

PUBMED 8626273

REFERENCE 9

AUTHORS Kim, S.R. and Komano, T.

TITLE The plasmid R64 thin pilus identified as a type IV pilus

JOURNAL J. Bacteriol. 179 (11), 3594-3603 (1997)

MEDLINE 97315231

PUBMED 9171405

REFERENCE 10

AUTHORS Narahara, K., Rahman, E., Furuya, N. and Komano, T.

TITLE Requirement of a limited segment of the sog gene for plasmid R64 conjugation

JOURNAL Plasmid 38 (1), 1-11 (1997)

MEDLINE 97428559

PUBMED 9281491

REFERENCE 11

AUTHORS Furuya, N. and Komano, T.

TITLE Mutational analysis of the R64 orit region: requirement for precise location of the Nika-binding sequence

JOURNAL J. Bacteriol. 179 (23), 7291-7297 (1997)

MEDLINE 98053841

PUBMED 9393692

REFERENCE 12

AUTHORS Yoshida, T., Furuya, N., Ishikura, M., Isobe, T., Haino-Fukushima, K., Ogawa, T. and Komano, T.

TITLE Purification and characterization of thin pili of IncII plasmids ColIB-P9 and R64: formation of Pili-specific cell aggregates by type IV pili

JOURNAL J. Bacteriol. 180 (11), 2842-2848 (1998)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 26, 2004, 15:20:44 ; Search time 347.756 Seconds
(without alignments)
3115.905 Million cell updates/sec

Title: US-10-676-299-9

Perfect score: 25

Sequence: 1 ttaagtcataatgtttttgactta 25

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_htg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_om.*

21: em_or.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pl.*

26: em_ro.*

27: em_sts.*

28: em_un.*

29: em_vi.*

30: em_htg_hum.*

31: em_htg_inv.*

32: em_htg_other.*

33: em_htg_mus.*

34: em_htg_pln.*

35: em_htg_rod.*

36: em_htg_mam.*

37: em_htg_vrt.*

38: em_sy.*

39: em_htgo_hum.*

40: em_htgo_mus.*

41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	25	100.0	3492	1	ECARSBC	X80057 E.coli gene
2	25	100.0	10240	1	AE000426	AE000426 Escherich
3	25	100.0	11524	1	AE015361	AE015361 Shigella
C 4	25	100.0	179941	2	AC145934	AC145934 Gallus ga
5	25	100.0	225419	1	EC00076	EC00076 E. coli chr
6	25	100.0	242495	2	AC146183	AC146183 Pan trogl
C 7	25	100.0	289816	1	AE016992	AE016992 Shigella
8	21.8	87.2	11071	1	AE005575	AE005575 Escherich
C 9	21.8	87.2	207991	2	EX545855	EX545855 Danio rer
10	21.8	87.2	267888	1	AP002565	AP002565 Escherich
C 11	20.8	83.2	71661	10	AL954643	AL954643 Mouse DNA
C 12	20.8	83.2	94879	9	HG384F21	AL022171 Human DNA
13	20.8	83.2	112484	10	AL954640	AL954640 Mouse DNA
14	20.8	83.2	151950	2	AC084813	AC084813 Homo sapi
C 15	20.2	80.8	2000	6	AX508952	AX508952 Sequence
16	20.2	80.8	5641	8	AX346299	AX346299 Sequence
17	20.2	80.8	89779	8	AB005234	AB005234 Arabidops
C 18	20.2	80.8	101371	8	AC002534	AC002534 Arabidops
19	20.2	80.8	146712	5	AL935136	AL935136 Zebrafish
C 20	20.2	80.8	182803	2	AC145899	AC145899 Pan trogl
21	20.2	80.8	185981	9	AC006459	AC006459 Homo sapi
C 22	20.2	80.8	216010	2	AC119317	AC119317 Rattus no
23	20.2	80.8	219476	9	AC068945	AC068945 Homo sapi
C 24	20.2	80.8	220414	2	EX546482	EX546482 Danio rer
C 25	20.2	80.8	240663	2	AC097541	AC097541 Rattus no
26	19.8	79.2	884	11	CNS06K7W	AL403338 T3 end of
27	19.8	79.2	57502	2	AC103790	AC103790 Homo sapi
C 28	19.8	79.2	68181	2	AC103792	AC103792 Homo sapi
C 29	19.8	79.2	94609	8	AB046439	AB046439 Arabidops
C 30	19.8	79.2	112615	9	AC079617	AC079617 Homo sapi
31	19.8	79.2	115626	9	AC103792	AC103792 Homo sapi
32	19.8	79.2	116461	8	AC051625	AC051625 Genomic S
C 33	19.8	79.2	121038	8	AP003342	AP003342 Oryza sat
34	19.8	79.2	121501	8	AC069557	AC069557 Genomic S
C 35	19.8	79.2	149098	8	AP003296	AP003296 Oryza sat
36	19.8	79.2	152813	2	AP005860	AP005860 Oryza sat
C 37	19.8	79.2	157295	2	AC079055	AC079055 Homo sapi
38	19.8	79.2	159302	2	AC098585	AC098585 Homo sapi
C 39	19.8	79.2	160295	10	AL845504	AL845504 Mouse DNA
C 40	19.8	79.2	160586	2	AC126319	AC126319 Mus muscu
C 41	19.8	79.2	176959	9	AC026402	AC026402 Homo sapi
C 42	19.8	79.2	182554	2	AC138307	AC138307 Mus muscu
C 43	19.8	79.2	183629	10	AC132596	AC132596 Mus muscu
44	19.8	79.2	189245	2	AC068090	AC068090 Homo sapi
45	19.8	79.2	195494	2	EX545912	EX545912 Danio rer

ALIGNMENTS

RESULT 1
ECARSBC
LOCUS ECARSBC 3492 bp DNA linear BCT 20-JUL-1995
DEFINITION E.coli genes arsR, arsB, arsC.
ACCESSION X80057
VERSION X80057.1 GI:510824
KEYWORDS arsB gene; arsC gene; arsenate reductase; arsenic-efflux pump;
arsenic-inducible repressor; arsR gene.
SOURCE Escherichia coli
ORGANISM Escherichia coli
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.
REFERENCE 1
AUTHORS Diorio,C., Cai,J., Marmor,J., Shinder,R. and DuBow,M.S.
TITLE An Escherichia coli chromosomal ars operon homolog is functional in

```

JOURNAL      arsenic detoxification and is conserved in gram-negative bacteria
MEDLINE      J. Bacteriol. 177 (8), 2050-2056 (1995)
PUBMED       95238276
REFERENCE    7721697
AUTHORS      2 (bases 1 to 3492)
              Diorio,C.
TITLE        Direct Submission
JOURNAL      Submitted (06-JUL-1994) C. Diorio, McGill University, 3775
              University Street, Montreal, Quebec, H3A 2B4, CANADA
FEATURES     Location/Qualifiers
             1..3492
               /organism="Escherichia coli"
               /mol_type="Genomic DNA"
               /strain="K12"
               /db_xref="taxon:562"
             repeat_region
               528..538
               /rpt_type=INVERTED
             repeat_region
               544..554
               /rpt_type=INVERTED
             gene
               655..1061
               /gene="arsR"
             -35_signal
               655..660
               /gene="arsR"
             -10_signal
               678..683
               /gene="arsR"
             RBS
               697..701
               /gene="arsR"
             CDS
               708..1061
               /gene="arsR"
               /codon_start=1
               /transl_table=11
               /product="arsenic-inducible repressor"
               /protein_id="CAA56361.1"
               /db_xref="GI:510825"
               /db_xref="GOA:P37309"
               /db_xref="SWISS-PROT:P37309"
               /translation="MSFLPTIQFLKILADETRLIGVLLSELGELCVCDLCTALDQSQ
               PRISRLHALLRESGLLDRKQKWHYRLSPHPAWAAKIIDEAWRCEQKVAIVRN
               LARQNGSGDSKNICS"
             gene
               1071..2404
               /gene="arsB"
             -35_signal
               1071..1076
               /gene="arsB"
             repeat_region
               1072..1084
               /rpt_type=INVERTED
             repeat_region
               1087..1099
               /rpt_type=INVERTED
             -10_signal
               1088..1093
               /gene="arsB"
             RBS
               1106..1110
               /gene="arsB"
             CDS
               1115..2404
               /gene="arsB"
               /codon_start=1
               /transl_table=11
               /product="membrane-located arsenite efflux pump"
               /protein_id="CAA56362.1"
               /db_xref="GI:516211"
               /db_xref="GOA:P37310"
               /db_xref="SWISS-PROT:P37310"
               /translation="MLLAGAIFVLITVLVIWQPKGLIGKSATLGAVALVLTGWHVHP
               DIPVWNVNNAFAFVILIIILLDESQFFWAALHVRMGNGRGLLFTWVILVLG
               AVALFANDGAALIPFIVIMLLALGSPKGTILFAVMAAGTAFTADTASPLIVSNLV
               NIVSDAFGLGFRYASVWPFVDIAALVATLVMHLHIFRKDIQNTYDMALLKSPAAEI
               KDPATQIVGVLVLLLVGVFLEPLGIPVSAIANGVALILFVVAKRGHAINTKVLR
               GAPQIVIFSLGMVTVYGLRNAGLTGLIPVSAIANGVALILFVVAKRGHAINTKVLR
               MNNPTVLVGLASIDGSTASGVKEAMVYANVIGCDLGPKITPTIGSLATLLHLVL
               KMTNISGVYFRTGIIWTLPLVPLVTLAALRLSFTL"
             gene
               2417..2842
               /gene="arsC"
             CDS
               2417..2842
               /gene="arsC"
               /codon_start=1

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 26, 2004, 15:20:44 ; Search time 347.756 Seconds
(without alignments)
3115.905 Million cell updates/sec

Title: US-10-676-299-10
Perfect score: 25
Sequence: 1 taagtcaaaacatatatgacttaa 25

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.roi.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg.hum.*
- 31: em.htg.inv.*
- 32: em.htg.other.*
- 33: em.htg.mus.*
- 34: em.htg.pln.*
- 35: em.htg.rod.*
- 36: em.htg.mam.*
- 37: em.htg.vrt.*
- 38: em.sy.*
- 39: em.htgo.hum.*
- 40: em.htgo.mus.*
- 41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
C 1	25	100.0	3492	1	ECARSRBC
C 2	25	100.0	10240	1	AE000426
C 3	25	100.0	11524	1	AE015361
C 4	25	100.0	179941	2	AC145934
C 5	25	100.0	225419	1	ECOUW76
C 6	25	100.0	242495	2	AC146183
C 7	25	100.0	289816	1	AE016992
C 8	21.8	87.2	11071	1	AE005575
C 9	21.8	87.2	207991	2	EX545855
C 10	21.8	87.2	267888	1	AP002565
C 11	20.8	83.2	11661	10	AL954643
C 12	20.8	83.2	96879	9	HS384F21
C 13	20.8	83.2	112434	10	AL954640
C 14	20.8	83.2	151950	2	AC084813
C 15	20.2	80.8	2000	6	AX508952
C 16	20.2	80.8	5641	6	AX348299
C 17	20.2	80.8	89779	8	AE005234
C 18	20.2	80.8	101371	8	AE002534
C 19	20.2	80.8	146712	5	AL133136
C 20	20.2	80.8	182803	2	AC146899
C 21	20.2	80.8	186981	9	AC006459
C 22	20.2	80.8	216010	2	AC119517
C 23	20.2	80.8	219476	9	AC068945
C 24	20.2	80.8	220414	2	EX546482
C 25	20.2	80.8	240663	2	AC097541
C 26	19.8	79.2	884	11	CNS06K7W
C 27	19.8	79.2	57902	2	AC103790
C 28	19.8	79.2	68181	2	AC103972
C 29	19.8	79.2	94609	8	AE046439
C 30	19.8	79.2	112615	9	AC079617
C 31	19.8	79.2	115626	9	AC103792
C 32	19.8	79.2	116461	8	AC051625
C 33	19.8	79.2	121038	8	AP003342
C 34	19.8	79.2	121501	8	AC069557
C 35	19.8	79.2	149098	8	AP003296
C 36	19.8	79.2	152813	2	AP005860
C 37	19.8	79.2	157295	2	AC079055
C 38	19.8	79.2	159302	2	AC098585
C 39	19.8	79.2	160295	10	AL845504
C 40	19.8	79.2	160586	2	AC126319
C 41	19.8	79.2	176959	9	AC026402
C 42	19.8	79.2	182554	2	AC138307
C 43	19.8	79.2	183629	10	AC132596
C 44	19.8	79.2	189245	2	AC068090
C 45	19.8	79.2	195494	2	EX545912

ALIGNMENTS

RESULT 1	ECARSRBC	3492 bp	DNA	linear	BCT 20-JUL-1995
LOCUS	E.coli genes arsr, arsb, arsc.				
DEFINITION	X80057.1	GI:510824			
ACCESSION	arSB gene; arsc gene; arsenate reductase; arsenic-efflux pump;				
VERSION	arSB gene; arsc gene; arsenate reductase; arsenic-efflux pump;				
KEYWORDS	arsenic-inducible repressor; arsr gene.				
SOURCE	Escherichia coli				
ORGANISM	Escherichia coli				
REFERENCE	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;				
AUTHORS	Enterobacteriaceae; Escherichia.				
TITLE	Diorio, C., Cai, J., Marmor, J., Shinder, R. and DuBow, M.S.				
	An Escherichia coli chromosomal ars operon homolog is functional in				

```
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
arsenic detoxification and is conserved in gram-negative bacteria
J. Bacteriol. 177 (8), 2050-2056 (1995)
75238276
7721697
2 (bases 1 to 3492)
Diorio C.
Direct Submission
Submitted (06-JUL-1994) C. Diorio, McGill University, 3775
University Street, Montreal, Quebec, H3A 2B4, CANADA
Location/Qualifiers
1..3492
/organism="Escherichia coli"
/mol_type="genomic DNA"
/strain="K12"
/db_xref="taxon:562"
repeat_region
528..538
/rpt_type=INVERTED
repeat_region
544..554
/rpt_type=INVERTED
gene
655..1061
/gene="arsR"
-35_signal
655..660
/gene="arsR"
-10_signal
678..683
/gene="arsR"
697..701
/gene="arsR"
708..1061
/gene="arsR"
/codon_start=1
/transl_table=1
/product="arsenic-inducible repressor"
/protein_id="CAA56361.1"
/db_xref="GI:510825"
/db_xref="GOA:P37309"
/db_xref="SWISS-PROT:P37309"
/translation="MSFLLPIOLFILADETRGLGIVLLSELGELCVCLTALDOSQ
PKIRHALLRESGLLDKQKQWVHSLPSHIPAWAKIIDEAWRCEQEKQVAIVEN
LARQCSGDSKNICS"
1071..2404
/gene="arsB"
-35_signal
1071..1076
/gene="arsB"
repeat_region
1072..1084
/rpt_type=INVERTED
repeat_region
1087..1099
/rpt_type=INVERTED
-10_signal
1088..1093
/gene="arsB"
RBS
1106..1110
/gene="arsB"
CDS
1115..2404
/gene="arsB"
/codon_start=1
/transl_table=1
/product="membrane-located arsenite efflux pump"
/protein_id="CAA56362.1"
/db_xref="GI:516211"
/db_xref="GOA:P37310"
/db_xref="SWISS-PROT:P37310"
/translation="MLLAGALFVLTIVIVIQKGLIGNSATLGAVALVTVGHPG
DIPVWVWVNTAFTAVTIIISLLDESGFFEAALHVSNGRGLLFTWVILIG
AAVALFANDAGAILIPVIAMLLALGFGKGTFLAFVMAAGFIATASPLIVSLV
NVSADFFGJGFREYASVMPVDIAAIVATVLMHLYFRKDI PNTYDMLALKSPARAI
KDPATFTGWWVLLVGLVGFVLEPLGPVSAIAAGVALILFVYAKRGHAINTKVLR
GAPQIVIFSLGMLVYVGLRNAGLTYELSGVNLVADGLMAATLGTGPTLFLSSI
MNNMPTLVGLASLDIGSTAGVINEAMVYANVIGCDLGPKITPGISLATLLWLHVSQ
KNWISGYVFTGIINTLIPVLFVTLAALALRSFTL"
2417..2842
/gene="arsC"
-35_signal
2417..2842
/gene="arsC"
/codon_start=1
transl_table=1
/product="arsenate reductase"
/protein_id="CAA56363.1"
/db_xref="GI:516212"
/db_xref="GOA:P37311"
/db_xref="SWISS-PROT:P37311"
/translation="NSNITVYHNPACGTSRNTLEIRNSGTPTTIHYLETPTTRDEL
VKLIADMGISVRALLRKKNVPEYELGLAEDKFDRLDLIDMLQHLINRPIVVTPLG
TLCRPSVLEILPDAQGAFSGKEDGKVDGAKRLK"
2922..2905
/rpt_type=INVERTED
repeat_region
2910..2923
/rpt_type=INVERTED
ORIGIN
Query Match 100.0%; Score 25; DB 1; Length 3492;
Best Local Similarity 100.0%; Pred No. 14;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TAAGTCAAAAACATATGACTTAA 25
Db 662 TAAGTCAAAAACATATGACTTAA 638
RESULT 2
AE0000426/c
LOCUS
DEFINITION
Escherichia coli K12 MG1655 section 316 of 400 of the complete
genome.
ACCESSION AE0000426 U00096
VERSION AE0000426.1 GI:1789910
KEYWORDS
SOURCE
Escherichia coli K12
ORGANISM
Escherichia coli K12
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.
REFERENCE
1 (bases 1 to 10240)
Blattner, F.R., Plunkett, G. III, Bloch, C.A., Perna, N.T., Burland, V.,
Riley, M., Soliardo-Vides, J., Glasner, J.D., Rode, C.K., Mayhew, G.F.,
Gregor, J.J., Davis, N.W., Kirkpatrick, H.A., Goeden, M.A., Rose, D.J.,
Mau, B. and Shao, Y.
The complete genome sequence of Escherichia coli K-12
Science 277 (5331), 1453-1474 (1997)
97426617
PUBMED
9278503
REFERENCE
2 (bases 1 to 10240)
Blattner, F.R.
Direct Submission
Submitted (16-JAN-1997) Guy Plunkett III, Laboratory of Genetics,
University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA.
Email: ecolig@genetics.wisc.edu Phone: 608-262-2534 Fax:
608-263-7459
3 (bases 1 to 10240)
Blattner, F.R.
Direct Submission
Submitted (03-SEP-1997) Guy Plunkett III, Laboratory of Genetics,
University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA.
Email: ecolig@genetics.wisc.edu Phone: 608-262-2534 Fax:
608-263-7459
4 (bases 1 to 10240)
Plunkett, G. III.
Direct Submission
Submitted (13-OCT-1998) Laboratory of Genetics, University of
Wisconsin, 445 Henry Mall, Madison, WI 53706, USA
This sequence was determined by the E. coli Genome Project at the
University of Wisconsin-Madison (Frederick R. Blattner, director).
Supported by NIH grants HG00301 and HG01428 (from the Human Genome
Project and NCHGR). The entire sequence was independently
determined from E. coli K12 strain MG1655. Predicted open reading
frames were determined using GeneMark software, kindly supplied by
Mark Borodovsky, Georgia Institute of Technology Atlanta, GA,
30332 [e-mail: mark@amber.gatech.edu]. Open reading frames that
have been correlated with genetic loci are being annotated with CG
```